*Inorg. Chem.* **2005**, *44*, 4923–4933

# XAFS Spectral Analysis of the Cadmium Coordination Geometry in Cadmium Thiolate Clusters in Metallothionein

**Jayna Chan,[†] Maureen E. Merrifield,[†] Alexander V. Soldatov,[‡] and Martin J. Stillman*[,†]**

*Department of Chemistry, The University of Western Ontario, London, Ontario N6A 5B7 Canada, and Faculty of Physics, Rostov State University, Sorge 5, Rostov-on-Don, 344090 Russia*

Received August 17, 2004

We report the combination of measurement and prediction of X-ray absorption fine structure (XAFS) data, where the term XAFS refers to the overall spectrum that encompasses both the X-ray Absorption Near Edge Structure (XANES) region as well as the Extended X-ray Absorption Fine Structure (EXAFS) region, to evaluate the cadmium thiolate cluster structures in the metalloprotein metallothionein. XAFS spectra were simulated using coordinates from molecular models of the protein calculated by molecular mechanics/molecular dynamics (MM3/MD), from NMR analyses, and from analysis of X-ray diffraction data. XAFS spectra were also simulated using the coordinates from X-ray crystallographic data for $[Cd(SPh)_4]^{2-}$, CdS, $[Cd_2(\mu\text{-}SPh)_2(SPh)_4]^{2-}$, and $[Cd_4(\mu\text{-}SPh)_6(SPh)_4]^{2-}$. The simulated XAFS data that were calculated using the FEFF8 program closely resemble the experimental data reported for $[Cd(SPh)_4]^{2-}$, CdS, $[Cd_2(\mu\text{-}SPh)_2(SPh)_4]^{2-}$, $[Cd_4(\mu\text{-}SPh)_6(SPh)_4]^{2-}$, rabbit liver metallothionein cadmium $\alpha$-domain ($Cd_4$-$\alpha$ MT), and cadmium rabbit liver $\beta\alpha$ metallothionein ($Cd_7$-$\beta\alpha$ MT). MM3 force field parameters were modified to include cadmium–sulfur bonding and were initially set to values derived from published X-ray diffraction and EXAFS experimental data. The force field was further calibrated and adjusted through comparison between experimental spectra taken from the literature and simulated XAFS spectra calculated using the FEFF8 program in combination with atomic coordinates from MM3/MD energy minimization. MM3/MD techniques were used with the calibrated force field to predict the high-resolution structure of the metal clusters in rabbit liver $Cd_7$-MT. Structures for $Cd_3S_9$ ($\beta$) MT and $Cd_4S_{11}$ ($\alpha$) MT domains from MM3/MD calculations and those previously reported for $Cd_7$-MT on the basis of $^1$H and $^{113}$Cd NMR data were compared. Structural differences between the different models for these cadmium thiolate clusters were evident. Combining the measurement and simulation of XAFS data provides an excellent method of assessing, modeling, and predicting metal-binding sites in metalloproteins when X-ray absorption spectroscopy (XAS) data are available.

## Introduction

Metals are estimated to play a role in 30% of protein molecules; indeed metal-dependent biological chemistries are vital for all forms of life. In most cases, it is the coordinated metal rather than the free ion that is the active species. However, determining the metal coordination environment in metalloproteins is a very difficult, yet vitally important task, because the biological function can be associated with the structural properties of the binding site. Obtaining high-resolution structures is necessary to understand the chemical reactivities of these metallobiomolecules more precisely. Reactions that occur at the molecular level in biological systems are highly dependent on the spatial proximity and orientation of the reacting species and adjacent amino acids side chains.

In many cases, spectroscopic data, typically from NMR, X-ray absorption spectroscopy (XAS), UV−visible absorption, and circular dichroism techniques, do not unambiguously identify the coordinating ligands of the bound metals. Even the geometry of the binding site is very difficult to determine from purely spectroscopic techniques. While X-ray diffraction data provide wide-scale detail of the whole biological molecule, the initial resolution may not be enough to unambiguously define the coordination geometry of a bound metal atom, especially in proteins not previously well-studied by other techniques. This problem becomes more

* Author to whom correspondence should be addressed. E-mail: Martin.Stillman@uwo.ca.
[†] University of Western Ontario.
[‡] Rostov State University. E-mail: soldatov@rsu.ru.

10.1021/ic048871n CCC: $30.25 
Published on Web 06/11/2005

acute when there are multiple metals in the binding site where even the number of metal atoms in the binding site might be difficult to determine precisely.

Synchrotron-based techniques such as XAS[1,2] have long provided invaluable structural parameters in terms of high accuracy bond lengths from both solid and solution samples of various metalloproteins.[3−14] X-ray absorption near edge structure (XANES) and extended X-ray absorption fine structure (EXAFS) spectroscopies are subsets of XAS that provide complementary structural information. The term X-ray absorption fine structure (XAFS) refers to the overall spectrum that encompasses both the XANES region as well as the EXAFS region, which we will use throughout this paper. XANES spectroscopy is sensitive to angular and nearest-neighbor effects, and EXAFS spectroscopy is sensitive to bond distances and coordination number. However, even with these techniques, identification of the coordination geometry cannot always be determined with high enough precision to identify the type, number, and geometry of all the attached ligands. This is especially a problem when the metal binding site is complicated, which is the case for metalloproteins with multiple metals in either one site or a number of isolated sites within the same molecule, because the EXAFS and XANES data are an average of all metal−ligand bonds.

Unambiguous confirmation of the structure of the metal binding site in the metalloprotein has proven difficult without X-ray diffraction data. And in some cases, particularly metallothionein, crystallization has proven to be difficult or impossible so that X-ray diffraction data cannot be obtained. Thus, new ways of improving the accuracy and precision of the determination of the coordination shell of the bound metals are needed for the structural elucidation of these metal-binding biomolecules.

Computational techniques offer a possible route to this goal by combining information from a number of experimental sources; however, proving the accuracy of computationally determined structures requires high-resolution experimental data. Molecular geometries of the metal binding site(s) cores determined using molecular mechanics/molecular dynamics (MM/MD),[15] semiempirical (i.e. MOPAC),[16] and ab initio (i.e. DFT)[17] techniques can also provide good estimates; however, a major concern is the effect of the peptide on the metal coordination geometry. To achieve this, the modeling calculations must include as much of the surrounding peptide environment as much as possible, which greatly increases the number of atoms included in the calculation. MM/MD has been widely used as the computational method of choice for biomolecules because it is capable of handling large molecules such as the metalloproteins within a computationally short time frame. Determining the accuracy of these fast MM/MD calculations has been extremely difficult.

A solution to the experimental problem of obtaining structures from X-ray diffraction is provided by XAFS spectral simulation. Development of efficient codes for the calculation of photoelectron multiple-scattering resonances and the self-consistent calculation of local potential functions[1,18−20] has recently allowed analysis of the spectroscopic features in XANES data. Multiple-scattering XANES calculations are more challenging than EXAFS calculations, since electron scattering is much stronger and a multiple-scattering regime for the photoelectron takes place at low energies and, hence, XANES data are much more sensitive to details of the scattering potential and to the distribution of atoms around the photoabsorbing ion.[2] Recently, a new method based on ab initio DFT simulations followed by XANES analysis was shown to provide powerful predictors of the structure of $Na_4Cl_4$ clusters.[21] Geometry-sensitive, metal-edge XAFS data from biologically bound metals can be directly compared with simulated XAFS data calculated for model compounds of the metal binding sites. For multiple metals in biological molecules, the XAFS data can be calculated for the environment proposed for each metal site individually and, as well, different ligand combinations can be tested in the coordination sphere. Through the comparison between the measured XAFS data from chemically synthesized analogue molecules and biomolecules, and the calculated XAFS data from theoretically generated models, the coordination environment that best matches the metal binding site in the biomolecule can be identified.

The metalloprotein metallothionein (MT) (Figure 1) is a biological metal chelator that has demonstrated high affinity


(1) Bianconi, A. In *X-ray Absorption: Principles, Applications and Techniques of EXAFS, SEXAFS and XANES*; Koningsberger, D. C., Prins, R., Eds.; John Wiley & Sons: New York, 1989; pp 573−662.
(2) Penner-Hahn, J. E. *Comprehensive Coordination Chemistry II*; Elsevier Ltd.: Oxford, U.K., 2004; Vol. 2.
(3) Smolentsev, G.; Soldatov, A. V.; Wasinger, E. C.; Solomon, E. I. *Inorg. Chem.* **2004**, *43*, 1825−1827.
(4) DellaLonga, S.; Pin, S.; Cortes, R.; Soldatov, A.; Alpert, B. *Biophys. J.* **1998**, *75*, 3154−3162.
(5) Boffi, F.; Ascone, I.; Longa, S. D.; Girasole, M.; Yalovega, G.; Soldatov, A. V.; Varoli-Piazza, A.; Castellano, A. C. *Eur. Biophys. J.* **2003**, *32*, 329−341.
(6) Parsons, J. G.; Aldrich, M. V.; Gardea-Torresdey, J. L. *Appl. Spectrosc. Rev.* **2002**, *37*, 187−222.
(7) Hasnain, S. S.; Garner, C. D. *Prog. Biophys. Mol. Biol.* **1987**, *50*, 47−65.
(8) Hasnain, S. S. *Life Chem. Rep.* **1987**, *4*, 273−331.
(9) George, G. N.; Hedman, B.; Hodgson, K. O. *Nat. Struct. Biol.* **1998**, *5*, 645−647.
(10) Garner, C. D.; Collison, D.; Pidcock, E. *Philos. Trans. R. Soc. London, Ser. A* **1996**, *354*, 325−357.
(11) Garner, C. D. *Physica B* **1995**, *208 & 209*, 714−716.
(12) Valentine, J. S. *Handb. Met.-Ligand Interactions Biol. Fluids: Bioinorg. Chem.* **1995**, *1*.
(13) Solomon, E. I.; Randall, D. W.; Glaser, T. *Coord. Chem. Rev.* **2000**, *200−202*, 595−632.
(14) Clark-Baldwin, K.; Tierney, D. L.; Govindaswamy, N.; Gruff, E. S.; Kim, C.; Berg, J.; Koch, S. A.; Penner-Hahn, J. E. *J. Am. Chem. Soc.* **1998**, *120*, 8401−8409.
(15) Leach, A. R. *Molecular modelling: principles and applications*, 2nd ed.; Prentice Hall: New York, Harlow, England, 2001.
(16) Stewart, J. J. *J. Comput. Aided Mol. Des.* **1990**, *4*, 1−105.
(17) Bickelhaupt, F. M.; Baerends, E. J. *Rev. Comput. Chem.* **2000**, *15*, 1−86.
(18) Ankudinov, A. L.; Bouldin, C. E.; Rehr, J. J.; Sims, J.; Hung, H. *Phys. Rev. B* **2002**, *65*, 104−107.
(19) Ankudinov, A. L.; Ravel, B.; Rehr, J. J.; Conradson, S. D. *Phys. Rev. B* **1998**, *58*, 7565−7576.
(20) Benfatto, M.; Congiu-Castellano, A.; Daniele, A.; Longa, S. D. *J. Synchrotron Radiat.* **2001**, *8*, 267−269.
(21) Yalovega, G.; Soldatov, A. V.; Riedler, M.; Pederson, M. R.; Kolmakov, A.; Nowak, C.; Moller, T. *Chem. Phys. Lett.* **2002**, *356*, 23−28.

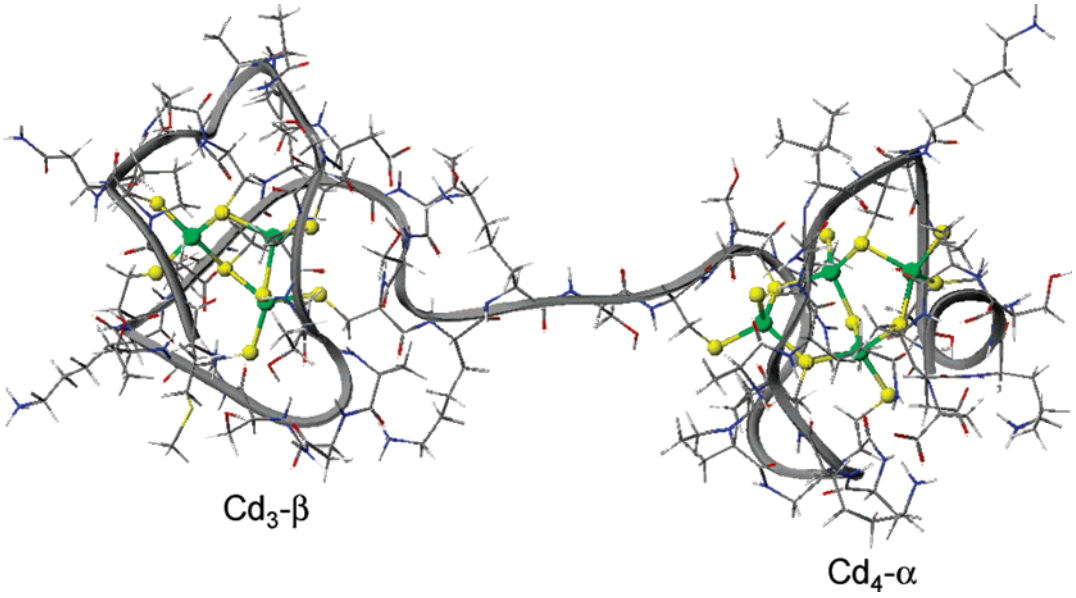


**Figure 1.** Ribbon structure of rabbit MT-2a showing the cadmium thiolate cluster connectivity based on MM3/MD calculations described in this paper using a force field modified from our previous work.[38] This figure illustrates the tight wrapping of the protein backbone around each cluster. The 20 cysteinyl sulfurs (yellow) form both bridging and terminal ligands for 7 bound cadmiums (green). The amino acid sequence of rabbit MT-2A is MDPNCSCAAAGDSCTCANSCTCKACKCTSCKKSCCSCCPPGCAKCAQGCI CKGASDKCSCCA.[24]

toward heavy metal ions and in mammals exhibits a 2-domain structure[22] when metalated. The structural properties of MT have been probed by several experimental techniques, including NMR spectroscopy,[23−27] X-ray crystallography,[25] and X-ray absorption spectroscopy (XAS),[28−33] as well as theoretical techniques.[34−43] Metal coordination in MT is known only to occur via cysteinyl sulfurs ligands with stoichiometries of $M_3S_9$ (in the $\beta$ domain) and $M_4S_{11}$ (in the $\alpha$ domain) for M(II) metals such as Cd(II) and Zn(II)[31] and $M_6S_9$ and $M_6S_{11}$ in the respective domains for M(I) metals such as Cu(I).[39,44] On the basis of experimental results, Cd is known to bind with tetrahedral coordination in these clusters, with a mixture of bridging and terminal sulfurs. Mass spectrometric data show that there are no counterions involved with the cadmium, and both charges are accommodated for by the thiolates of the cysteines.[45]

The connectivities between the bridging and terminal sulfurs were initially determined by NMR spectroscopy and X-ray crystallography.[25,26] The resolution of the model structures generated from analysis of the NMR data is not very high with regards to the metal−ligand bond length or precise geometry at the metal. There are unusually strained geometries at these metal binding sites. Indeed, it is not clear if these strained sites are necessary components of the structure or artifacts of the model-building process. Naturally built-in strain would suggest that the metal sites were not relaxed and, therefore, metal transfer to a site of lower energy would take place readily. Clearly, a relaxed, lowest energy set of sites might suggest that the metal binding is more permanent. Thus, the exact structure becomes important in determining possible function.

Previous reports of modeling metallothionein have included calculations of the metal thiolate core by ab initio methods,[37] MM/MD of the complete protein[38−40,42,46] and

(22) Boulanger, Y.; Goodman, C. M.; Forte, C. P.; Fesik, S. W.; Armitage, I. M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 1501−1505.
(23) Messerle, B. A.; Schaeffer, A.; Vasak, M.; Kagi, J. H. R.; Wuthrich, K. *J. Mol. Biol.* **1990**, *214*, 765.
(24) Arseniev, A.; Schultze, P.; Worgotter, E.; Braun, W.; Wagner, G.; Vasak, M.; Kagi, J. H.; Wuthrich, K. *J. Mol. Biol.* **1988**, *201*, 637.
(25) Braun, W.; Vasak, M.; Robbins, A. H.; Stout, C. D.; Wagner, G.; Kagi, J. H.; Wuthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10124−10128.
(26) Otvos, J. D.; Armitage, I. M. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 7094−7098.
(27) Schultze, P.; Worgotter, E.; Braun, W.; Wagner, G.; Vasak, M.; Kagi, J. H.; Wuthrich, K. *J. Mol. Biol.* **1988**, *203*, 251−268.
(28) Jiang, D. T.; Heald, S. M.; Sham, T. K.; Stillman, M. J. *J. Am. Chem. Soc.* **1994**, *116*, 11004−11013.
(29) Gui, Z.; Green, A. R.; Kasrai, M.; Bancroft, G. M.; Stillman, M. J. *Inorg. Chem.* **1996**, *35*, 6520−6529.
(30) Garner, C. D.; Hasnain, S. S.; Bremner, I.; Bordas, J. *J. Inorg. Biochem.* **1982**, *16*, 253−256.
(31) Stillman, M. J. *Coord. Chem. Rev.* **1995**, *144*, 461−511.
(32) George, G. N.; Byrd, J.; Winge, D. R. *J. Biochem.* **1988**, *263*, 8199−8203.
(33) George, G. N.; Winge, D.; Stout, C. D.; Cramer, S. P. *J. Inorg. Biochem.* **1986**, *27*, 213−220.
(34) Chang, C.-C.; Huang, P. C. *Protein Eng.* **1996**, *9*, 1165−1172.
(35) Chang, C.-C.; Liao, W.-F.; Huang, P. C. *Protein Eng.* **1998**, *11*, 41−46.
(36) Chang, C. C.; Huang, P. C. In *Metallothionein IV*; Klaassen, C., Ed.; Birkhauser Verlag: Basel, Switzerland, 1999; pp 45−49.
(37) Enescu, M.; Renault, J.-P.; Pommeret, S.; Mialocq, J.-C.; Pin, S. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3762−3767.
(38) Fowle, D. A.; Stillman, M. J. *J. Biomol. Struct. Dyn.* **1997**, *14*, 393−406.
(39) Presta, A.; Fowle, D. A.; Stillman, M. J. *J. Chem. Soc., Dalton Trans.* **1997**, *6*, 977−984.
(40) Chan, J.; Huang, Z.; Merrifield, M. E.; Salgado, M. T.; Stillman, M. J. *Coord. Chem. Rev.* **2002**, *233−234*, 319−339.
(41) Berweger, C. D. *Proteins: Struct., Funct., Genet.* **2000**, *41*, 299−315.
(42) Romero, N.; Capdevila, M.; Gonzalez-Duarte, P.; Oliva, B. *J. Mol. Model.* **1996**, *2*, 417−426.
(43) Szilagyi, Z.; Fenselau, C. *Drug Metab. Dispos.* **2000**, *28*, 174−179.
(44) Bertini, I.; Hartmann, H.-J.; Klein, T.; Liu, G.; Luchinat, C.; Weser, U. *Eur. J. Biochem.* **2000**, *267*, 1008−1018.
(45) Blanc, J. C. Y. L.; Presta, A.; Veinot, J.; Gibson, D.; Siu, K. W. M.; Stillman, M. J. *Protein Pept. Lett.* **1997**, *4*, 313−320.

protein-drug complexes,[43] and some cases of restraining the amino acids of the peptide framework in their original positions but leaving the cysteines and coordinated metals free for semiempirical calculations.[34,35,41] However, in each case, it was not possible to judge the accuracy of the calculated model.

The purpose of this paper is to determine the high-resolution metal binding site geometry of the cadmium thiolate clusters in MT by analysis of XAFS data. We demonstrate that the combination of classical molecular mechanics/molecular dynamics (MM/MD) modeling of the metal binding site and the subatomic resolution provided by XAFS spectroscopy is a powerful approach for studying structure−function relationships in many important biochemical compounds. We show how XAFS data calculated for theoretically constructed model structures can be used as a sensitive means of verifying the structural accuracy in metal−protein binding sites by comparison with experimental XAFS data. The energy range of all the spectra shown is 26 700−26 900 eV, which includes the XANES region plus part of the EXAFS region.

## Materials and Methods

A $[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$ sample was provided by Prof. P. A. W. Dean at The University of Western Ontario[47] in crystalline form and was used without further purification. The data were measured on the bending magnet beam line 20 (PNC-CAT) of the Advanced Photon Source (Chicago, IL). The data were measured in the transmission yield mode and processed using the WinXAS 97 v. 2.0 program. The samples were finely ground and applied to Mylar tape. The tape was folded 32 times to add to the density of the powder sample; the layering of the Mylar itself does not affect the spectrum. This method of sample preparation also greatly reduces the chance of there being any pinholes. We do not expect any reaction with the aluminum in the aluminized Mylar with the thiolates because any bare aluminum surface immediately forms a very stable oxide layer.

XAFS data for a number of complexes and protein samples were generously provided by Prof. G. George and Prof. S. S. Hasnain and are redrawn here from the digital data. XAFS data for $[Cd(SPh)_4]^{2-}$, CdS, and $[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$ were provided by George from the work of Pickering et al.[48] XAFS data for $[Cd(SPh)_4]^{2-}$, $[Cd_2(\mu\text{-SPh})_2(SPh)_4]^{2-}$, $[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$, rabbit liver $Cd_7$-MT ($\beta\alpha$-MT), and rabbit liver $Cd_4Co_3$-MT are unpublished spectra provided by Hasnain.[49] Data for CdS were obtained from The XAFS Model Compound Library[50] posted by CARS, The University of Chicago, APS ANL at http://cars9.uchicago.edu/~newville/ModelLib/. Experimental XAFS data for rabbit liver $Cd_7$-MT ($\beta\alpha$-MT) were reprocessed from experimental files used by Jiang et al.[28]

**Software used.** The CAChe Worksystem Pro version 6.1.1 molecular modeling software (Fujitsu America) was used for the molecular mechanics (MM3) energy minimization with augmented force fields for the cadmium-thiolate bonds described below. The molecular dynamics (MD) was carried out using the CAChe Worksystem Pro version 6.1.1 molecular modeling software (Fujitsu America) with the MM3 system for energy minimization. The FEFF8.2 software package[18,19] was used to calculate the simulated XAFS data. The WinXAS 97 v. 2.0 program was used to process the experimental data measured as a part of this work.

**Structures of the Small Thiolate Complexes of Cadmium.** The structures of $[Cd(SPh)_4]^{2-}$,[51] $[Cd_2(\mu\text{-SPh})_2(SPh)_4]^{2-}$,[52] and $[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$ [53] were taken from the Cambridge Crystallographic Database. The structure of $[Cd_2(\mu\text{-SPh})(SPh)_6]^{3-}$ was constructed and minimized using the CAChe Worksystem Pro version 6.1.1 molecular modeling software (Fujitsu America). The atomic coordinates of CdS and $[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$ were obtained from the Cambridge Structural Database,[54] and the atomic coordinates of the cadmium-containing $\beta$ and $\alpha$ domains of rabbit (2MRB, 1MRB)[24] metallothionein and the $\alpha$ domain of rat metallothionein (4MT2)[25] were obtained from the RCSB protein databank[55] posted at http://www.pdb.org/. All structures were rendered using the CAChe Worksystem Pro version 6.1.1 molecular modeling software (Fujitsu America).

**Force Field Modifications in MM3 Calculations.** The original MM3 force field reported by Allinger,[56−58] was modified to include parameters for cadmium−sulfur interactions. The modifications included the addition of a cadmium atom type and defining bond lengths for the terminal Cd−$S_t$ (2.46 Å) and bridging Cd−$S_b$ (2.56 Å) thiolates [initially based on the X-ray structure of $[Cd_4(SC_6H_5)_{10}]^{2-}$ by Hagen et al.[53] and EXAFS data for metallothionein[28,29] and then later fine-tuned], tetrahedral bond angles S−Cd−S (109.5°) and Cd−S−Cd (109.5°), and bent bond angle Cd−S−C (95.9°) (for both terminal and bridging sulfurs).

**Molecular Dynamics Calculations.** MD calculations were carried out on the molecular models by applying the following sequence of calculations to each molecular structure: (i) MM3 energy minimization using the modified force field; (ii) MD simulation for 500 ps at a set temperature of 300 K using the modified force field; (iii) MM3 repeated on the lowest energy conformer; (iv) the cycle repeated from ii to iii until no further energy minima were observed in the MD trajectory. This is the same procedure used previously for minimization of the energies of the metallothionein structures with zinc, cadmium, mercury, and copper.[38,39] The MM3 minimized structure calculated at this point was used for the XAFS data simulation.

**XAFS Data Simulations.** For each molecular structure obtained from protein database files, crystallographic database files, and the MM3/MD calculations described above, atoms within a radius of 5.5 Å surrounding each absorbing cadmium (which typically includes about 40 atoms) were extracted prior to XAFS simulation. This, therefore, includes not just the first shell atoms (as is the typical use of FEFF8.2) but additional shells of atoms that have been shown to contribute to the overall XAFS spectrum. The atomic


(46) Stillman, M. J.; Thomas, D.; Trevithick, C.; Guo, X.; Siu, M. *J. Inorg. Biochem.* **2000**, *79*, 11−19.
(47) Dean, P. A. W.; Vittal, J. J. In *Metallothioneins*; Stillman, M. J., III, C. F. S., Suzuki, K. T., Eds.; VCH: New York, 1992; pp 346−386.
(48) Pickering, I. J.; Prince, R. C.; George, G. N.; Rauser, W. E.; Wickramasinghe, W. A.; Watson, A. A.; Dameron, C. T.; Dance, I. G.; Fairlie, D. P.; Salt, D. E. *Biochim. Biophys. Acta* **1999**, *1429*, 351−364.
(49) Hasnain, S. S. Unpublished.
(50) Newville, M.; Carroll, S.; O'Day, P.; Waychunas, G. http://cars9.uchicago.edu/~newville/ModelLib/.
(51) Swenson, D.; Baenziger, N. C.; Coucouvanis, D. *J. Am. Chem. Soc.* **1978**, *100*, 1932−1934.
(52) Abrahams, I. L.; Garner, C. D.; Clegg, W. *J. Chem. Soc., Dalton Trans.* **1987**, 1577−1579.
(53) Hagen, K. S.; Holm, R. H. *Inorg. Chem.* **1983**, *22*, 3171−3174.
(54) Allen, F. H. *Acta Crystallogr.* **2002**, *B58*, 380−388.
(55) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(56) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566.
(57) Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566−8575.
(58) Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8576−8582.

coordinates of these atoms were used to simulate the XAFS spectra for each metal site of each molecule using the FEFF8.2 software package.[18,19] These ab initio calculations are based on self-consistent, final-state potentials. The scattering potential is approximated by overlapped spherical muffin-tin potentials (including a core hole); overlapping the muffin tins by about 15% compensates partly for the errors in the spherical approximation. Many-body effects are incorporated approximately in terms of a complex, energy-dependent self-energy, which replaces the exchange-correlation potential of ground-state density functional theory. We used the Dirac−Hara energy-dependent potential with the addition of the imaginary part of the Hedin−Lundqvist electron-gas self-energy. These self-energy terms naturally add important final-state broadening and self-energy shifts but do not contribute any new features to the spectra.

## Results and Discussion

The focus of our interest is specifically the Cd−SR clusters that model the Cd−$S_{cys}$ cluster structures in metallothionein. The chemically reasonable model structures used here are shown in Figure 2, which are representative of the metal coordination environments in metallothionein. We have simulated the Cd K-edge XAFS data for each structure and compare those simulated XAFS data with the experimental data to obtain the higher precision required for the determination of the properties of the metal-binding site coordination geometry. This also provides a means of assessing the accuracy of models on the basis of analysis of NMR spectra and X-ray diffraction data and the results of MM/MD calculations. We describe the individual XAFS data below.

**Structures of the Cd−S Complexes.** Figure 2A,B represents the $\beta$ and α metal thiolate clusters for the rabbit and rat isoforms of metallothionein. We indicate the number of the cysteine residue in the peptide chain (1−9 in $\beta$ and 1−11 in α). The metal thiolate cluster models used here to explore the thiolate function of cysteine include $[Cd(SPh)_4]^{2-}$ (Figure 2C), which demonstrates simple Cd−S bonding character, $[Cd_2(\mu\text{-SPh})(SPh)_6]^{3-}$ (Figure 2D), which was chosen to introduce the simplest bridging and terminal sulfur effects (XAFS data not shown), $[Cd_2(\mu\text{-SPh})_2(SPh)_4]^{2-}$ (Figure 2E), which was chosen to illustrate the effects of strain, as well as a more bridged structure; the adamantane-like cluster ($[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$)[53,59,60] (Figure 2F), which is a good model of the α domain cluster ($Cd_4S_{11}$) (Figure 2B), including both bridging and terminal sulfurs, and finally the crystalline CdS (Figure 2G), which is a highly symmetric array of bridged thiolates.

Of these model compounds, $[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$ best represents the α-domain structure. The α-domain and the adamantane Cd−S clusters involve four cadmium atoms tetrahedrally coordinated via terminal and bridging cysteinyl sulfurs, the difference being two terminal sulfurs (in the α domain of MT) in place of a single bridged sulfur in the adamantane structure. On the other hand, the $\beta$-domain is simpler, with two terminal thiolates for each of the three cadmiums, and may be best represented by the $[Cd_2(\mu\text{-SPh})_2(SPh)_4]^{2-}$ model.

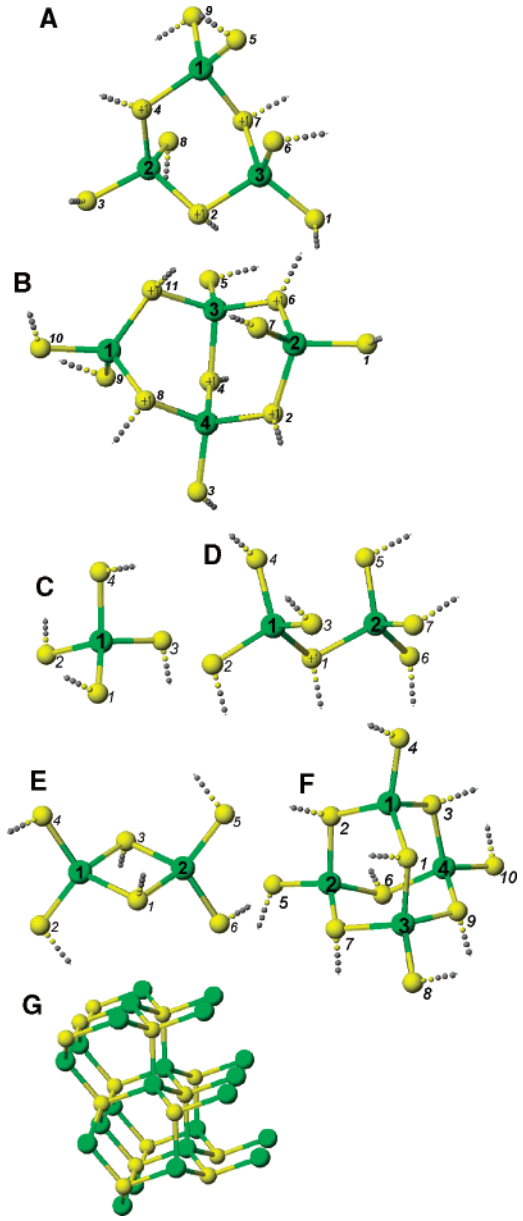


**Figure 2.** Cadmium thiolate complexes: (A) $Cd_3S_{cys9}$ $\beta$-domain of MT; (B) $Cd_4S_{cys11}$ α-domain of MT; (C) $[Cd(SPh)_4]^{2-}$; (D) $[Cd_2(\mu_2\text{–SPh})(SPh)_6]^{3-}$; (E) $[Cd_2(\mu_2\text{-SPh})_2(SPh)_4]^{2-}$; (F) $[Cd_4(\mu_2\text{−SPh})_6(SPh)_4]^{2-}$; (G) CdS. Structures A, B, and D were drawn from MM3/MD models calculated in this work. Structures C,[51] E,[52] and F[53] were drawn from crystallographic data deposited in the Cambridge Crystallographic Database. Each Cd atom is tetrahedrally coordinated by sulfur atoms. The dotted lines represent extended bonding to the cysteine ligands (for metallothionein) and phenyl rings (for the simple cadmium thiolate complexes). Key: yellow = sulfur; green = cadmium(II); gray = carbon.

**Energy Minimization by MM3/MD.** Comparison between the simulated XAFS data for the MM3/MD energy-minimized structures described in Figure 2 and available experimental data provides an excellent method for identifying the coordination environments in metalloproteins. The force field was adjusted by comparing the simulated XAFS data with experimental data for the small thiolate complexes. Differences between the original structures and the energy-minimized structures may be accounted for by a combination of effects but most likely (i) crystal packing vs vacuum, (ii)

(59) Dean, P. A. W.; Vittal, J. J. *Inorg. Chem.* **1986**, *25*, 514−519.
(60) Lee, G. S. H.; Fisher, K. J.; Vassallo, A. M.; Hanna, J. V.; Dance, I. G. *Inorg. Chem.* **1993**, *32*, 66−72.

poor experimental resolution, (iii) experimental sample contamination, or (iv) improperly parametrized force field.

For the small molecules, MM3 alone provides a very good result and molecular dynamics results in little change to the energy-minimized structure. However, for the peptides, the role of molecular dynamics is essential because the structures are dependent on so many interrelated connections extending across the molecular framework, including those far from the metal binding site. We have previously noted that MD calculations over the first 100 ps result in a much more relaxed structure than obtained from MM3 alone. Over the next 200−5000 ps, the metal thiolate core is seen to be flexible, with no systematic change in structural alignment, even when the MD is carried out at very high simulated temperatures (1000 K).[40] In this extended period, the main structural fluctuations are almost entirely concerned with minor changes in alignment of parts of the peptide chain that are not closely connected with the metal binding site. Therefore, there is no single identifiable source of strain that contributes to the strains imposed on the $Cd_3S_9$ ($\beta$) and $Cd_4S_{11}$ ($\alpha$) units; rather, there is an accumulation of both short- and long-range effects.

The MM3/MD method must account for the conformations of the peptide chain, as well as accommodate the strain imposed by the tight wrapping of the peptide chain as it forms the clustered cadmium-binding site. The three-dimensional structure of metalated metallothioneins is largely controlled by the connectivities of the cysteinyl sulfurs (Figures 1 and 2A,B). We expect the effect of the peptide chain compared with simple thiolate ligands to be seen in distortions in the bond lengths and angles away from the pure tetrahedral geometries (but not in excess). The presence of both terminal and bridging thiolates will also lead to distortions in the structure. These distortions will raise the energies of specific sites, changing the metalation/demetalation chemistry, which may be closely connected with the function of metallothionein.

The metal−cysteine connectivities of metallothionein were established by analysis of NMR data.[26] These deposited MT structures,[23,24] however, do show obvious distortions from tetrahedral geometry at more than one site, as well as uncharacteristic bond lengths and angles (described later in further detail). Highlighted (by a double underscore) in Supporting Information Table 1 are distortions in the NMR structures identified as values that lie outside range of Cd−S and S−Cd−S literature values as determined from the Cambridge Structural Database.[54] Values that are at the limits of published data are indicated by a single underscore. We do not reproduce these distortions in the MM3/MD energy minimized models of the NMR structure of metallothionein. These energy-minimized structures instead showed relaxation of the bond lengths and bond angles at each of the metal sites of the cluster in metallothionein reported here and a consequential improvement in the spectroscopic properties when compared with the experimental data (Figure 4).

**Comparison between the Simulated and Experimental XAFS Spectra for $[Cd(SPh)_4]^{2-}$, $[Cd_2(\mu_2\text{-}SPh)(SPh)_6]^{3-}$, $[Cd_2(\mu_2\text{-}SPh)_2(SPh)_4]^{2-}$, and $[Cd_4(\mu_2\text{-}SPh)_6(SPh)_4]^{2-}$.** Figure 3 shows experimental and simulated Cd K-edge XAFS data for the complexes shown in Figure 2 (C−G), with the exception of $[Cd_2(\mu_2\text{-}SPh)(SPh)_6]^{3-}$. For each of the complexes shown, we compare experimental data and simulated data calculated on the basis of the X-ray structures. In addition, for $[Cd_2(\mu_2\text{-}SPh)_2(SPh)_4]^{2-}$ and $[Cd_4(\mu_2\text{-}SPh)_6(SPh)_4]^{2-}$, we have energy minimized the structures and have

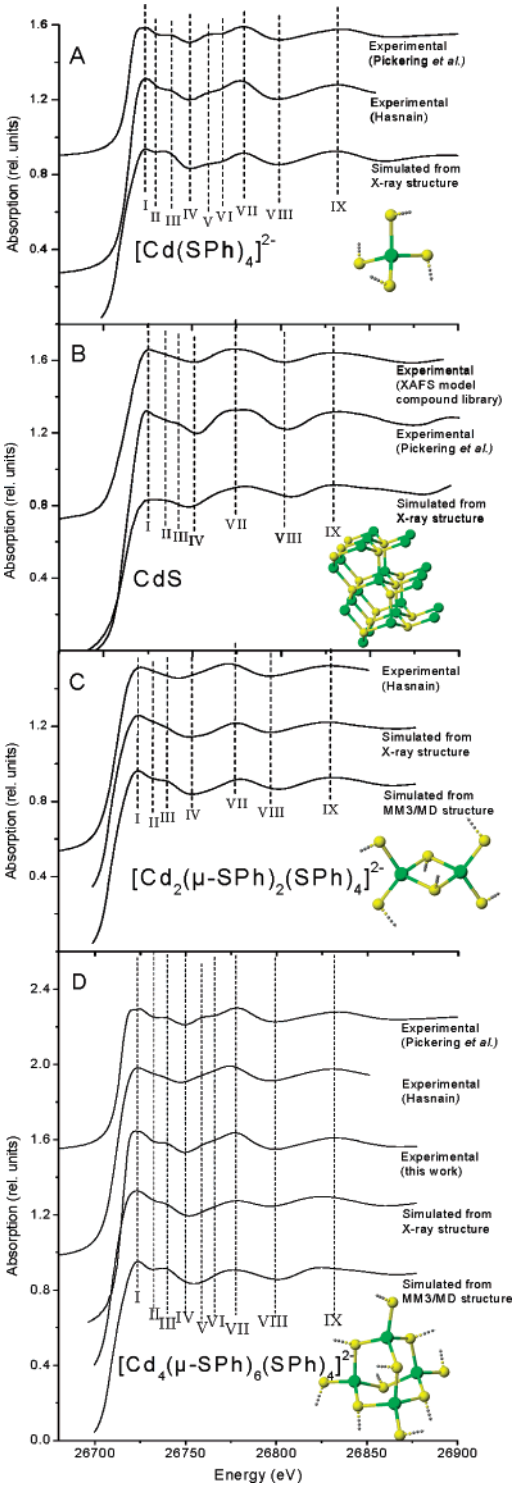


**Figure 3.** Comparison of experimental XAFS spectra with XAFS spectra simulated from coordinates on the basis of X-ray structures and MM3/MD calculations: (A) $[Cd(SPh)_4]^{2-}$;[48] (B) CdS;[49] (C) $[Cd_2(\mu\text{-}SPh)(SPh)_6]^{3-}$; (D) $[Cd_2(\mu\text{-}SPh)_2(SPh)_4]^{2-}$; (E) $[Cd_4(\mu\text{-}SPh)_6(SPh)_4]^{2-}$.

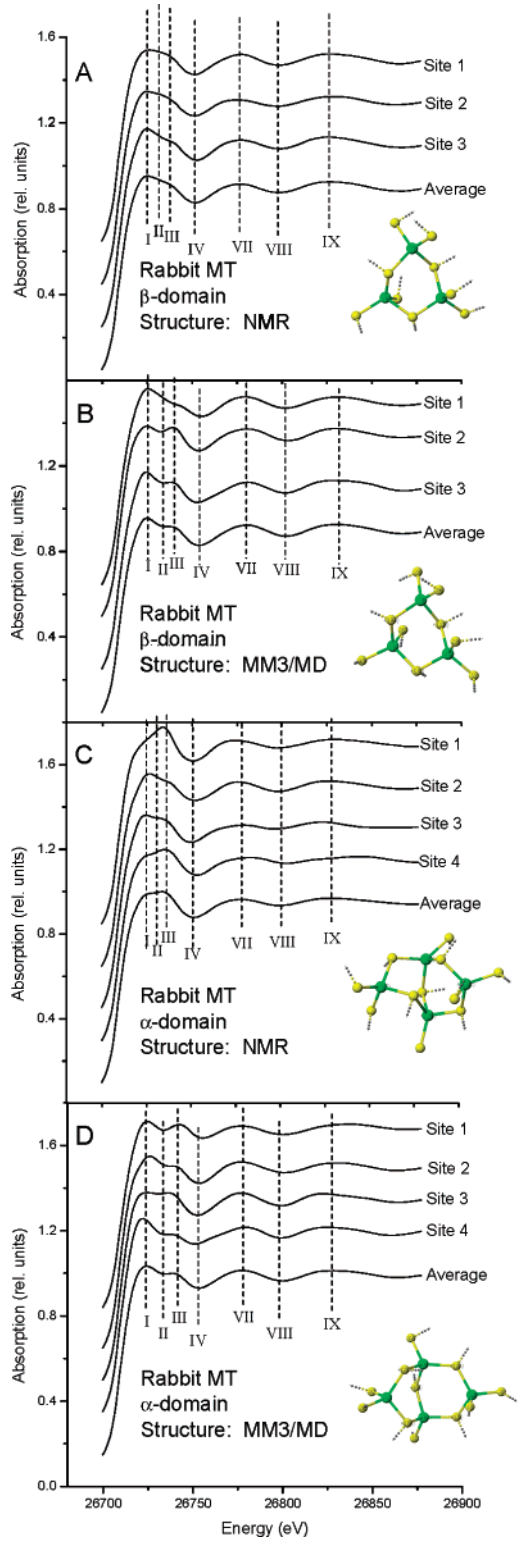


**Figure 4.** Site-specific and cluster-averaged XAFS spectra for the $\beta$ and $\alpha$ domains of rabbit liver MT for structures prior to and following energy minimization. The numbering scheme of the features shown here follows the numbering scheme in Figure 3.

simulated the XAFS using these atomic coordinates. In Figure 3B−D the simulated XAFS data were calculated for each metal site in the cluster and the average was calculated. The spectra for the individual sites provide information on the contributions of specific sites to the overall spectrum, whereas the averages are what would be measured experimentally.

There are up to 9 characteristic features observed in each of the experimental and simulated spectra labeled I−IX, where features I, III, V, VII, and IX represent peak maxima and features II, IV, VI, and VIII represent peak minima. Features I−VI lie within the XANES region, and features VII−IX lie within the EXAFS region. These features are labeled on all subsequent spectra to aid in the comparison of spectral properties.

**$[Cd(SPh)_4]^{2-}$ Data.** Pickering et al.[48] and Hasnain[49] have measured experimental Cd K-edge XANES region spectra for $Cd(SPh)_4^{2-}$, which are shown together in Figure 3A. The X-ray diffraction coordinates of this structure[51] (Figure 2C) were used to calculate the simulated spectrum shown in Figure 3A. Each of the major features align closely, although the feature labeled III is more pronounced in the simulated spectrum in comparison to the experimental spectra. Overall, the simulation correctly reproduces the major features in the experimental data.

**CdS Data.** Two sets of experimental XAFS data are shown for CdS greenockite crystal, (Figure 2G), reproduced from the XAFS model compound library[50] and redrawn from the original data of Pickering et al.[48] Figure 3B. The simulated XAFS spectrum for CdS was calculated using a single site within the bulk crystal. The experimental data shown here appear to be less resolved than those in $[Cd(SPh)_4]^{2-}$. The 5 distinct features in these spectra align with features found in the $[Cd(SPh)_4]^{2-}$ spectra. The simulated XAFS data closely reproduce the structural effects (in number and periodicity) at the near-edge scattering region although the edge intensity of feature I is not as intense in the simulation, resulting in a lack of resolution at this energy.

**$[Cd_2(\mu_2\text{-SPh})(SPh)_6]^{3-}$ Data.** The structure (Figure 2D) consists of two metal sites where each site has 3 terminal and 1 bridging sulfur. This molecule could be thought of as two $[Cd(SPh)_4]^{2-}$ clusters bridged by a single sulfur, such that a spectrum similar to that for $[Cd(SPh)_4]^{2-}$ is anticipated. The simulated XAFS spectrum (not shown) based on MM3/MD calculations alone (because no X-ray diffraction data were available) closely resemble the data for $[Cd_2(\mu_2\text{-SPh})_2(SPh)_4]^{2-}$, with the exception of a rather more intense feature III.

**$[Cd_2(\mu_2\text{-SPh})_2(SPh)_4]^{2-}$ Data.** This structure (Figure 2E) was used to study the effect of strain on the XAFS spectra, because, although it is structurally similar to $[Cd_2(\mu_2\text{-SPh})(SPh)_6]^{3-}$, each cadmium site is now coordinated to 2 terminal and 2 bridging sulfurs. There is significant strain in the bonding as indicated by the deviation in $S_b-Cd-S_b$ bond angles away from tetrahedral geometry to approximately 90°.[52] The MM3/MD model also reproduces the geometric parameters of the X-ray structural data, with the exception of the alignment of the phenyl groups in the model. The model, which is calculated for an isolated molecule, cannot reproduce the effects of crystal packing on the conformations of the phenyl groups.

Figure 3C shows the experimental XAFS data from Hasnain,[49] and the simulation calculated from the X-ray

diffraction data coordinates. In addition we show the simulation calculated from the MM3/MD model coordinates. The overall features in the experimental spectrum are reproduced in the spectrum simulated from the X-ray diffraction data coordinates. The spectral data resemble those of $[Cd(SPh)_4]^{2-}$, though there are energy shifts observed in the features as well as some variations in broadening and peak intensity. In particular, there is a reduction in intensity of feature III as well as the shifting of features to lower energies in the EXAFS region. According to the Natoli rule[1] these energy shifts are a clear signal of the different interatomic distances between the photoabsorbing ion and the neighboring atoms, such that the longer the bond, the shorter the periodicity observed in the spectrum. These results satisfy the Natoli rule because the average $Cd-S_b$ lengths are 261 pm and the $Cd-S_t$ lengths are 248 pm, which compares with the $Cd-S_t$ lengths in $[Cd(SPh)_4]^{2-}$ of 253 pm. Similar results concerning the sensitivity of XAFS to the nearest ligand distances were obtained recently for transferrin.[5] It is not clear why there is a slight mismatch in energies of features IV and VII between the simulation and the experimental data, but this same effect is seen in Figure 3D. The simulation from the MM3/MD model also closely resembles both the experimental data and the simulation based on the X-ray diffraction data coordinates.

**$[Cd_4(\mu\text{-}SPh)_6(SPh)_4]^{2-}$ Data.** The structure of this complex (Figure 2F) contains four metal sites, which defines its adamantane-like character. This is the most similar structure to the α domain of metallothionein. The complex contains distorted $Cd_4$ tetrahedron bridged with 6 thiolates. The $CdS_4$ units are distorted tetrahedral with the connecting bridging thiolates giving $S_b-Cd-S_b$ bond angles with a range from 91.5 to 118°.[53] The terminal bonds are approximately 9 pm shorter than the bridging bonds. This structure was used to simulate the experimental XAFS data in Figure 3D. MM3/MD methods were used to model the structure resulting in bond lengths and bond angles that matched the X-ray data quite closely, but again the model represents an isolated molecule rather than the crystalline environment of the X-ray diffraction results.

Figure 3D shows the experimental and simulated XAFS spectra for this structure. Three sets of experimental data are shown: from Pickering et al.,[48] Hasnain,[49] and recorded as part of this work. XAFS simulations based on the X-ray structure and the MM3/MD model structure are also included. Correlation between the experimental and the simulated energies, amplitude, periodicity, and shape of the features is close. Although the simulation from the X-ray data matches the feature shapes for the data recorded for all the experimental data well, it is apparent that the average of the three data sets is the closest match. This suggests that measurement conditions and crystal packing effects between the three experiments is probably responsible for the differences in the feature III region. It is clear that the simulation from the model again exaggerates the intensity of feature III.

**Comparison of the Experimental and Simulated Data.** We find that there is excellent correlation between experimental and simulated XAFS data for $[Cd(SPh)_4]^{2-}$, $[Cd_2(\mu_2\text{-}SPh)_2(SPh)_4]^{2-}$, $[Cd_4(\mu_2\text{-}SPh)_6(SPh)_4]^{2-}$, and CdS (Figure 3), with all features in the experiment being reproduced in the simulations. The major mismatch is for feature III, which has more intensity in the calculation than we observe experimentally. Overall though, the comparisons between the experimental data and the simulations show that the FEFF8 method adequately reproduce the features of the XAFS spectra, especially in the XANES region. In addition to using simulations from X-ray diffraction data, the comparisons above included simulations based on MM3/MD models, which were designed to test the force field used in the MM3 calculations. Again the simulations closely match the experimental data.

It is clear from the X-ray structural data and the energy minimized models that the cadmium environment is not highly symmetric, with a wide range of values being found for both angles and bond lengths. The simulations allow study of individual sites in clustered molecules, which is the case for the $[Cd_4(\mu_2\text{-}SPh)_6(SPh)_4]^{2-}$ and the metallothionein cadmium thiolate clusters. As the complexities of the clusters increase, the shift away from perfect tetrahedral geometry increases, as shown by the increase in standard deviation observed in the bond lengths and bond angles of the small thiolate complexes (see the crystallographic data). The greater standard deviation at higher complexity signifies the presence of natural distortions. XAFS simulations using both the original deposited coordinates and the energy-minimized coordinates provide a means of examining the effects of these changes.

**Structures of the Metallothioneins.** While the structures of a number of cadmium (and, by inference, also zinc) metallothioneins have been determined using NMR techniques, only one structure has been reported from X-ray diffraction studies.[25] In both of these experimental techniques, the proposed model in the region of the metal thiolate complex is dominated by the effects of the protein backbone, which can lead to much less precision in these metal-based bond lengths and bond angles than in small molecules. Molecular modeling, using any energy minimization technique, can provide much higher resolution of these geometric parameters. The challenge is to assess the quality of each model for large molecules. XAFS provides a tool with which to achieve this because it is metal-based and specifically sensitive to the geometry of the metal. In this study we use simulation of the XAFS data based on the atomic coordinates for the metal core, compared with experimental data, to assess the accuracy of models of the metal binding site in metallothionein.

**XAFS Data for Rabbit Liver Metallothionein.** Structures of the separate β and α domains of rabbit liver (rlMT) cadmium metallothionein have been deposited in the protein databank from analyses of NMR data.[24] The simulated XAFS spectra for these structures are shown in Figure 4. Experimental XAFS data for $Cd_7$-rlMT from Jiang et al. and Hasnain[49] and for $Cd_4Co_3$-rlMT from Hasnain[49] are shown in Figure 5. We first examine the simulated data calculated from the deposited and minimized structures. The simulation

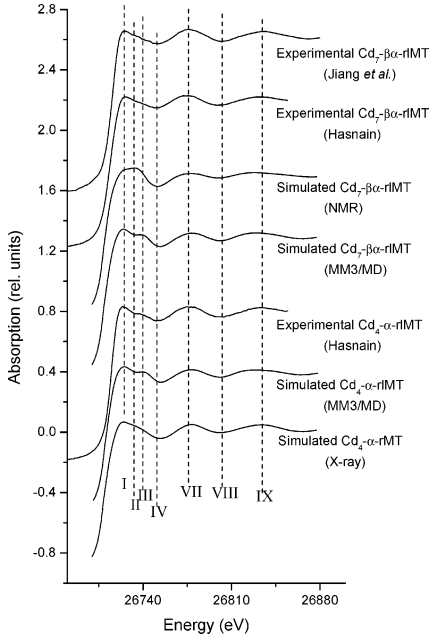


**Figure 5.** Comparison of experimental XAFS spectra for $Cd_7$-$\beta\alpha$-rlMT and $Cd_4$-$\alpha$-rlMT to XAFS spectra simulated using atomic coordinates from NMR, MM3/MD, and X-ray diffraction data.

procedure allows calculation of the contributions to the averaged XAFS spectrum from each of the individual cadmium sites (three in the $\beta$ domain and four in the $\alpha$ domain) prior to and following energy minimization. Variation between site-specific spectra is expected due to the lack of symmetry in the protein domain based on the different strains caused by the protein backbone, as well as the variability of the amino acid environment surrounding the cluster (included within the 5.5 Å shell).

In the $Cd_3$-$\beta$ rlMT NMR model structure, there is unusual geometry around metal site 2, which was identified by an overextended bond length of 270.4 pm that lies outside the realistic range (242−268 pm), as well as a rather acute bond angle of 89.4° that lies at the extreme of the realistic range (87.6−140°) (Supporting Information Table 1). The realistic ranges were determined by statistical analysis of all existing $CdS_4$ complexes deposited in the Cambridge Crystallographic Database.[54] XAFS spectroscopy allows us to determine the extent of impact that specific sites have on the overall spectrum. On the basis of the sensitivity of the XAFS spectra, we expect the spectrum for this unusual site to be significantly different from the spectrum generated from this site after energy minimization.

Shown in Figure 4A,B are the spectra calculated using the rlMT $\beta$-domain NMR structure prior to and following energy minimization, respectively. The simulated spectra based on the NMR coordinates (Figure 4A) exhibit the 9 characteristic features that appear to be attributable to $CdS_4$-type molecular structures as illustrated in Figure 3. There are distinctive dissimilarities between these site-specific spectra. The differences are observed as variations in the intensities of sites I−III, as well as minor energy shifts. The energy shift of interest is in feature VII of site 2, which is shifted to lower energy in comparison to sites 1 and 3 that is attributable to the overextended bond length. In this way, the simulated XAFS spectrum for each site allows assessment of the reasonableness of the geometric parameters that define that site. Of course, the experimental data show an average of all sites.

Energy minimization of $Cd_3$-$\beta$ rlMT results in significant changes in geometry at each of the metal sites (Supporting Information Table 1), resulting in more realistic bond lengths and bond angles. The simulated XAFS data, on the basis of these revised coordinates (Figure 4B), exhibit differences from the NMR-structure-based simulated spectrum (Figure 4A) that correlate with the changes in bond lengths and bond angles. Features II and III are much more resolved. However, the predicted contributions from individual sites still differ, as expected. The spectral changes observed for site 2 is more apparent than sites 1 and 3, which may be the result of the relaxation of site 2. Feature VII of site 2 is also realigned with sites 1 and 3, as expected from a more realistic bond length. As before for the small models, the simulations overemphasize the intensity of feature III. The variations between site-specific spectra demonstrate the nonequivalence of the 3 metal sites. The averaged spectrum for the minimized cluster is quite comparable to the spectra observed for the small molecules and to the experimental data shown in Figure 5. There is a very good match between the simulated XAFS spectrum for the energy-minimized $\beta$-domain and the experimental and simulated XAFS spectra for $[Cd_2(\mu_2\text{-}SPh)_2(SPh)_4]^{2-}$.

The simulated spectra based on the $\alpha$ domain of rlMT are shown in Figure 4C,D. Unusual geometries in the NMR structure are found at metal sites 3 (which has an overextended bond length of 272.4 pm and extreme angles of 92.2, 135.2, and 90.5°) and 4 (which has an overextended bond length of 269.5 pm and an extreme bond length of 243.7 pm) in this $\alpha$-domain (Supporting Informaton Table 1). The simulated XAFS spectra for each site on the basis of the deposited coordinates from the NMR analyses (Figure 4C) show dramatic variations in the intensity of features II and III. Sites 1 and 4 are unusual in that features II and III are higher in intensity than feature I. There are also energy shifts between spectra of the individual sites at features IV−VII, suggesting dramatic variation in bond lengths throughout the cluster. The averaged spectrum for this NMR-structure-based simulation is quite unlike the $CdS_4$-type spectra previously described above, which would imply that the adamantane-like cluster poorly models the $\alpha$-domain of MT.

Energy minimization of the $\alpha$-domain structure using the NMR structural coordinates as a starting point completely remove the extreme geometric parameters that are shown in Supporting Information Table 1. Figure 4D shows the simulated XAFS spectra for the four cadmium sites and the average spectrum on the basis of this energy-minimized structure. This figure shows that the individual site-specific spectra generated using the energy-minimized rlMT $\alpha$-domain structure are significantly different from one another, reflecting differences in the sites themselves. However, these differences are not so pronounced as in the data calculated directly from the NMR structural coordinates, Figure 4C. The energy shifts of feature VII better reflect the different

number of bridging and terminal bonds. The averaged XAFS spectrum for the energy-minimized $Cd_4$-α rlMT now shows close similarity to the $CdS_4$-type spectra (Figure 3D) and the experimental data shown in Figure 5, indicating that the adamantane-like cluster is a reasonable model for the α-domain of MT. The average of the combined simulated XAFS spectra for the two domains (Supporting Information Figure 1) provides the spectrum that would be recorded for the complete $Cd_7$-$\beta\alpha$ MT protein, which are shown in Figure 5.

**X-ray Diffraction Structure.** The only cadmium X-ray diffraction results to date for metallothionein have been from the rat liver (rMT) $Cd_5,Zn_2$-$\beta\alpha$MT.[25] The two Zn atoms are located in the $\beta$ domain so we have only calculated the XAFS spectra for the $Cd_4$-α domain structure deposited from the X-ray analysis. The X-ray model does not appear to have any unusual geometries (Supporting Information Table 1). The simulated XAFS for the X-ray structure of $Cd_4$-α rMT is shown in Figure 5G, with the individual contributions from each site shown as Supporting Information Figure 2A. There is significant similarity between the XAFS spectra for sites 2−4. The averaged simulated spectrum is similar to typical $CdS_4$-type molecules, though features II and III are not very distinctive. The energy-minimized α domain of rMT results in the simulated spectra shown in Supporting Information Figure 2B. There are significant changes in each of the site-specific spectra in comparison to the X-ray spectra, but the energy shifts of each of the features (and therefore the bond lengths) are comparable. Comparison of the averaged spectra for the two structures show a high degree of similarity, though features II and III are more well-defined in the spectrum of the minimized structure. As in all the simulated spectra, feature III is overemphasized. Not unexpectedly, the simulated XAFS data for the $Cd_4$-α-rlMT (Figure 5F) closely resembles that of the $Cd_4$-α-rMT (Supporting Information Figure 2B). These results allow comparisons between the simulated XAFS for solid-state structures (experimental X-ray diffraction data from a powder sample) and gas phase structures (energy-minimized). It is not unexpected that the cadmium thiolate clusters in the protein would be largely unchanged by the phase because these structures are buried within the protein with well-defined three-dimensional structures that dominate the overall structure of the protein.

**Comparison of Simulated XAFS Data on the Basis of Cadmium Thiolate Clusters in Metallothionein from NMR, X-ray Diffraction, and MM3/MD Coordinates with Experimental XAFS Data.** Figure 5 shows experimental XAFS data for $Cd_7$-rlMT from Jiang et al. (A) and Hasnain (B) and data for $Cd_4Co_3$-rlMT from Hasnain (E). In addition, we show the simulated XAFS spectra for the protein samples with geometries established by NMR (C), MM3/MD minimization (D, F), and X-ray diffraction (G). The experimental spectra for $Cd_7$-rlMT and the $Cd_4$-α-rlMT closely match each other (Figure 5A,B,E). The energies of the features in the $Cd_7$-rlMT simulated spectrum based directly on the NMR structure (C) do not align very well in the EXAFS region, and the shape of the XANES regions shows extremely poor correlation. The spectrum following energy minimization of the $Cd_7$-rlMT structure (D) shows much better alignment in the EXAFS region, and a much better agreement in the XANES region, with the exception of feature III, which we have commented previously, is overemphasized. It is clear that better data are required in the XANES region to more firmly establish the shape of the spectrum and the energies of these features.

The experimental data for $Cd_4$-α-MT (E) were obtained from data of Hasnain,[49] which included three cobalt atoms in the $\beta$-domain. These data can be used to assess the structural parameters for the α-domain based on X-ray coordinates. The simulated XAFS data using the atomic coordinates from the energy-minimized α-domain (F) and the X-ray diffraction results (G) are very similar and closely resemble the experimental data (E).

## Concluding Remarks

In the study reported here, we have reassessed the quality of the cadmium thiolate cores in two different metallothioneins by simulating the XAFS spectrum from the deposited structures, by energy-minimizing these structures, and by comparing the simulated data with experimental data from metallothionein and a number of models. Simulated and experimental data from small molecule models of the metallothionein binding sites allow for assessment of the calculations using FEFF8 methods of the simulated spectra as well. Calculating the XAFS data for each site in the $[Cd_4(\mu\text{-}SPh)_6(SPh)_4]^{2-}$ adamantane analogue and the $Cd_4S_{11}$ α-domain provides a much higher resolution assessment of similarities and differences because the XAFS experiment is sensitive to changes on a much smaller scale than other techniques.

These calculations have also allowed us to study the precision of the structures proposed on the basis of NMR and X-ray diffraction analyses using energy-minimized structures. The force field developed as part of this work shows that simulation of metal binding using MM3/MD techniques is possible and the results closely match the experimentally determined structures. These results are independent of the modeling technique; any modeling program will do (i.e. DFT, MOZYME, etc.), while these calculations may provide comparable or higher quality structures, using QM/MD for periods of 500 ps with 20 fs step sizes (as used in this work) requires 10 000 individual geometry minimizations, at a considerable cost in machine time. The comparison between the XAFS data from the MM3/MD models and the experimental data shows that the MM3/MD procedure is effective, while being fast enough to allow routine use.

This technique can be used to confirm that metal-based geometries calculated by separating the metal core from the protein result in reasonable geometric parameters. It can also greatly speed up the ability to accurately predict structures of newly discovered metallothioneins and provide insight into the search for functional aspects of structure and to questions such as "why are some metallothioneins one domain and some two?" Examples of these complications may be seen in the determination of the cadmium thiolate structures in

the lobster $Cd_6$-MT[61] and the seaweed *Fucus vesiculosus* $Cd_6$-MT.[62] The problem becomes more difficult, again considering just the metallothioneins, when metallothioneins containing other metals are studied. Optical data for a range of metallothioneins, including arsenic, copper, mercury, silver, gold, bismuth, and platinum, show that each of these species exist in solution; however, the presence of a number of metals in clusters (up to 18 for the 20-cysteine mammalian proteins) significantly complicates analysis of the data. In some cases (e.g. silver(I) replacement of copper(I)[63]) $^1$H NMR can provide information detailed enough for a model to be constructed, yet there have previously been few means of proving the correctness of the proposed structures. XAFS spectra are sensitive to all aspects of these multimetal clusters and can be measured relatively easily. Calculation of simulated XAFS spectra for each metal site and comparison with the proposed structure provides a means of verifying the accuracy of the model. A significant interest is the structural properties of copper-containing metallothioneins for which a structure has been proposed on the basis of the analysis of CD data.[39,64]

The results presented in this paper describe a technique with which to assess structural parameters of biological, metal-coordination structures proposed by analysis of spectroscopic data. For metal binding proteins, a major challenge is to identify the type, number, and geometry of coordinated ligands. The experimental requirement is to obtain XAFS spectra of the protein using the metal edge and as many small molecule models as possible and then model the metal binding site with the best available computational tools, calculate the simulated XAFS spectrum, and compare experimental data for the protein, models, and simulation. This greatly reduces the need to synthesize each of the model compounds.[48,65,66] In many cases, the XAFS spectra of the small molecules have been previously reported and X-ray diffraction data are also available. This clearly saves time and effort used to synthesize each compound that is required to prove the coordination properties of metals bound in proteins. The key requirement then is the availability of force fields that describe metal coordination geometry for use in the molecular mechanics calculations.

**Acknowledgment.** We wish to thank the Centre for Chemical Physics at the University of Western Ontario for the award of a Senior Visiting Fellowship (to A.V.S.). The research described here is part of a collaborative project between M.J.S. and Prof. P. R. Norton at The University of Western Ontario on biotechnological properties of metallothioneins. We thank the NSERC of Canada for financial support (to M.J.S.), RBFR for grant #05-04-49050 (to A.V.S.), the Ontario Graduate Scholarship program (J.C. and M.E.M.), Fujitsu America for providing advanced copies of the modeling software, Dr. G. Purvis (Fujitsu America) for technical help with the development of the force field for the bridging thiolates in the CAChe software, Prof. P. A. W. Dean for the gift of the $[Cd_4(\mu\text{-SPh})_6(SPh)_4]^{2-}$, Profs. R. R. Martin and T. K. Sham, and Dr. S. J. Naftel for assistance with measurement of the XAFS data on the bending magnet beam line 20 (PNC-CAT) at APS, Chicago, IL, and Dr. D.-T. Jiang for the XAFS measurements on $Cd_7$-MT. We thank Prof. S. S. Hasnain for providing the experimental XAFS data (unpublished) shown in Figures 3 and 5 and Prof. G. George for providing the experimental XAFS data shown in Figure 3.

**Supporting Information Available:** Figures showing contributions of the XAFS spectra of the individual domains to the XAFS spectrum of the complete molecule in rabbit liver metallothionein (Figure S1) and site-specific and averaged XAFS spectra for the α domain of rat MT prior to and following energy minimization (Figure S2) and a table listing connectivities and Cd−S bond lengths and angles in the $\beta$ and α domains of rabbit and rat metallothioneins (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

IC048871N

(61) Zhu, Z.; DeRose, E. F.; Mullen, G. P.; Petering, D. H.; Shaw, C. F. *Biochemistry* **1994**, *33*, 8858−8865.
(62) Morris, C. A.; Nicolaus, B.; Sampson, V.; Harwood: J. L.; Kille, P. *Biochem. J.* **1999**, *338*, 553−560.
(63) Peterson, C. W.; Narula, S. S.; Armitage, I. M. *FEBS Lett.* **1996**, *379*, 85−93.
(64) Presta, A.; Green, A. R.; Zelazowski, A.; Stillman, M. J. *Eur. J. Biochem.* **1995**, *227*, 226−240.
(65) Krebs, B.; Henkel, G. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 769−788.
(66) Santos, R. A.; Gruff, E. S.; Koch, S. A.; Harbison, G. S. *J. Am. Chem. Soc.* **1990**, *112*, 9257−9263.